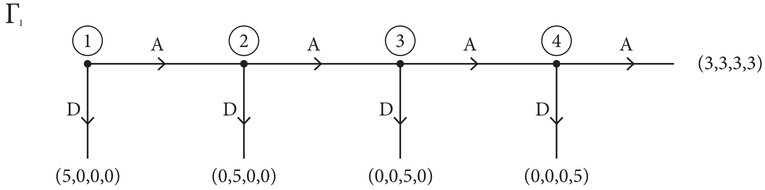
Time-consistent and Strategically Supported Cooperation in Dynamic Games

Leon A. Petrosyan

St. Petersburg University

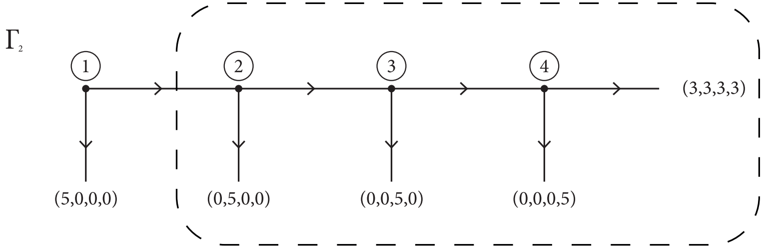
Example.



(D, D, D, D) – NE, (A, A, A, A) – NOT NE

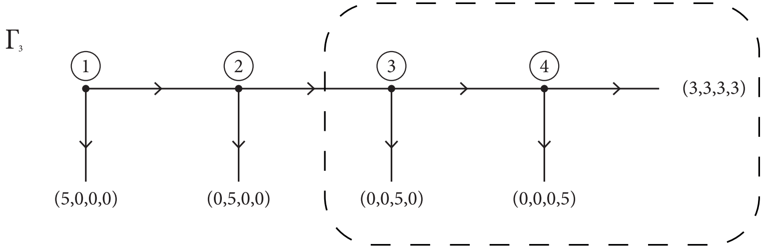
Characteristic Function of the game Γ_1 (C.f. of Γ_1)
$v_1(1,2,3,4) = 12$, $v_1(1,2,3) = 5$, $v_1(1,3,4) = 5$, $v_1(2,3,4) = 0$, $v_1(1,2,4) = 5$,
$v_1(1,2) = 5$, $v_1(1,3) = 5$, $v_1(1,4) = 5$, $v_1(2,3) = 0$, $v_1(2,4) = 0$, $v_1(3,4) = 0$,
$v_1(1) = 5$, $v_1(2) = 0$, $v_1(3) = 0$, $v_1(4) = 0$.

$Sh^1 = \left(\frac{27}{4}, \frac{7}{4}, \frac{7}{4}, \frac{7}{4}\right)$



C.f. of Γ_2
$v_2(1,2,3,4) = 12,\ v_2(1,2,3) = 5,\ v_2(1,3,4) = 5,\ v_2(2,3,4) = 9,$
$v_2(1,2) = 5,\ v_2(1,3) = 0,\ v_2(1,4) = 0,\ v_2(2,3) = 5,\ v_2(2,4) = 5,\ v_2(3,4) = 0,$
$v_2(1) = 0,\ v_2(2) = 5,\ v_2(3) = 0,\ v_2(4) = 0.$

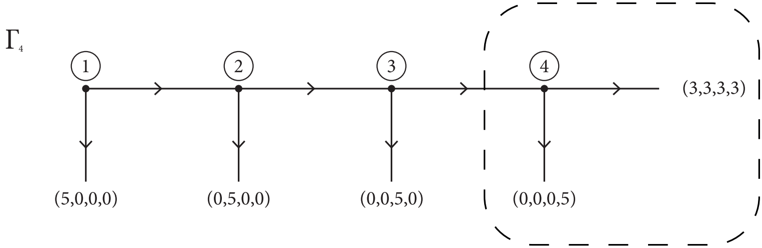
$Sh^2 = \left(\frac{19}{12}, \frac{65}{12}, \frac{30}{12}, \frac{30}{12}\right)$



C.f. of Γ_3

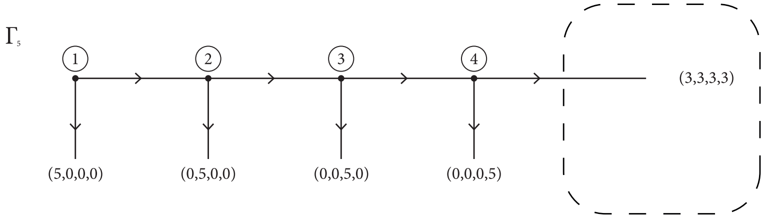
$v_3(1,2,3,4)=12,\ v_3(1,2,3)=5,\ v_3(1,3,4)=9,\ v_3(2,3,4)=9,\ v_3(1,2,4)=0$
$v_3(1,2)=0,\ v_3(1,3)=5,\ v_3(1,4)=0,\ v_3(2,3)=5,\ v_3(2,4)=0,\ v_3(3,4)=6,$
$v_3(1)=0,\ v_3(2)=0,\ v_3(3)=5,\ v_3(4)=0.$

$$Sh^3=\left(1,1,\tfrac{90}{12},\tfrac{30}{12}\right)$$



C.f. of Γ_4
$v_4(1,2,3,4) = 12,\ v_4(1,2,3) = 0,\ v_4(1,3,4) = 9,\ v_4(2,3,4) = 9,\ v_4(1,2,4) = 9$
$v_4(1,2) = 0,\ v_4(1,3) = 0,\ v_4(1,4) = 5,\ v_4(2,3) = 0,\ v_4(2,4) = 5,\ v_4(3,4) = 5,$
$v_4(1) = 0,\ v_4(2) = 0,\ v_4(3) = 0,\ v_4(4) = 5.$

$Sh^4 = \left(\frac{17}{12}, \frac{17}{12}, \frac{17}{12}, \frac{93}{12}\right)$



C.f. of Γ_5

$v_5(1,2,3,4) = 12,\ v_5(1,2,3) = v_5(1,3,4) = v_5(2,3,4) = v_5(1,2,4) = 9$
$v_5(1,2) = v_5(1,3) = v_5(1,4) = v_5(2,3) = v_5(2,4) = v_5(3,4) = 6,$
$v_5(1) = v_5(2) = v_5(3) = v_5(4) = 3.$

$Sh^5 = (3,3,3,3)$

IDP (Imputation Distribution Procedure)

$\beta_k, k = 1, \ldots, 5$

$Sh^1 = \beta_1 + Sh^2, Sh^2 = \beta_2 + Sh^3, \ldots, Sh^4 = \beta_4 + Sh^5$

$\beta_1 = (Sh^1 - Sh^2), \beta_2 = (Sh^2 - Sh^3), \beta_3 = (Sh^3 - Sh^4), \beta_4 = (Sh^4 - Sh^5), \beta_5 = Sh^5$

$\sum_{k=1}^{5} \beta_k = Sh^1, \sum_{k=2}^{5} \beta_k = Sh^2, \sum_{k=3}^{5} \beta_k = Sh^3,$
$\sum_{k=4}^{5} \beta_k = Sh^4, \sum_{k=5}^{5} \beta_k = Sh^5$

$\beta_1 = (\frac{62}{12}, -\frac{44}{12}, -\frac{9}{12}, -\frac{9}{12})$

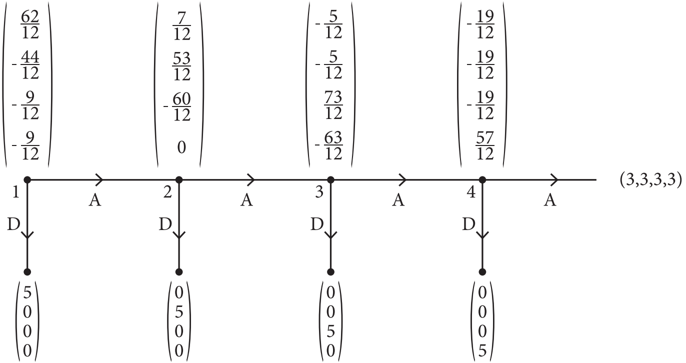
$\beta_2 = (\frac{7}{12}, \frac{53}{12}, -\frac{60}{12}, 0)$

$\beta_3 = (-\frac{5}{12}, -\frac{5}{12}, \frac{73}{12}, -\frac{63}{12})$

$\beta_4 = (-\frac{19}{12}, -\frac{19}{12}, -\frac{19}{12}, \frac{57}{12})$

$\beta_5 = (3, 3, 3, 3)$

Associated Game $\overline{\Gamma}$, and NE Strategically Supported Cooperation



(A, A, A, A) – NE

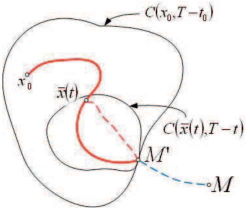
$$
NE \begin{cases} \frac{62}{12} + \frac{7}{12} - \frac{5}{12} - \frac{19}{12} + 3 > 5 \\ \frac{53}{12} - \frac{5}{12} - \frac{19}{12} + 3 > 5 \\ \frac{73}{12} - \frac{19}{12} + 3 > 5 \\ \frac{57}{12} + 3 > 5 \end{cases}
$$

1. Classical control problem.

$$\dot{x} = f(x, u), \quad x \in R^n, u \in U \subset CompR^l,$$
$$x(t_0) = x_0, \quad t \in [t_0, T],$$
$$H(x(T)) = -\rho(x(T), M).$$

$C(x_0, T - t_0)$ – reachability set.
$\bar{x}(t)$ – optimal trajectory.

$\Gamma(x_0, T - t_0), \quad \Gamma(\bar{x}(t), T - t), \quad C(\bar{x}(t), T - t)$



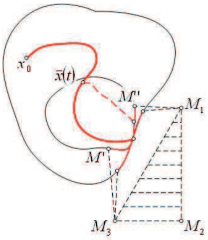
R. Bellmann
Time-consistency, Strong Time-consistency.

2. Multycriterial control.

$$\dot{x} = f(x, u), \quad x \in R^n, u \in U \subset CompR^l,$$
$$x(t_0) = x_0, \quad t \in [t_0, T],$$
$$H(x(T)) = \{H_1(x(T)), \ldots, H_k(x(T))\}.$$

Let $k = 3$, $H_i(x(T)) = -\rho(x(T), M_i)$.

Pareto-optimal solution.
$\bar{x}(t)$ – Pareto-optimal trajectory.
$\Gamma(x_0, T - t_0)$, $\Gamma(\bar{x}(t), T - t)$, $C(x_0, T - t_0)$, $C(\bar{x}(t), T - t)$, $P(x_0, T - t_0)$, $P(\bar{x}(t), T - t)$



TC but not STC

3. Nash bargaining solution in Differential Games.

$$\dot{x} = f(x, u_1, \dots, u_n), \quad x \in R^m, u_i \in U_i \subset CompR^l,$$
$$x(t_0) = x_0, \quad t \in [t_0, T],$$

The payoff of player $i \to H_i(x(T))$,
$\Gamma(x_0, T - t_0)$
$W(x_0, T - t_0; \{i\})$ – the garanteed payoff of player i

NB.

$$\max_{x' \in C(x_0, T - t_0)} \prod_{i=1}^{n} (H_i(x') - W(x_0, T - t_0; \{i\})) = \prod_{i=1}^{n} (H_i(\bar{\bar{x}}) - W(x_0, T - t_0; \{i\}))$$

$\bar{x}(t), \quad x_0 \to \bar{\bar{x}}, \quad \Gamma(\bar{x}(t), T - t), t \in [t_0, T - t_0], \quad W(\bar{x}(t), T - t; \{i\})$

$$\max_{x' \in C(\bar{x}(t), T - t)} \prod_{i=1}^{n} (H_i(x') - W(\bar{x}(t), T - t; \{i\})) = \prod_{i=1}^{n} (H_i(\bar{\bar{x}}(\bar{x}(t))) - W(\bar{x}(t), T - t; \{i\}))$$

$\bar{\bar{x}}(\bar{x}(t)) \neq const \neq \bar{x}$
NB, not TC, not STC

4. Differential Cooperative Game.

Differential Cooperative Game $\Gamma(x_0, T - t_0)$ with prescribed duration $T - t_0$ from the initial position x_0.

$$\dot{x} = f(x, u_1, \ldots, u_n), \quad x \in R^n, u_i \in U_i \tag{1}$$

integral payoff

$$K_i(x_0, T - t_0; u_1, \ldots, u_n) = \int_{t_0}^{T} h_i(x(t))dt, \quad h_i > 0, i = 1, \ldots, n.$$

Cooperative form of $\Gamma(x_0, T - t_0)$.

Cooperative behavior $u^*(t) = \{u_1^*(t), \ldots, u_n^*(t)\}$

$$\begin{aligned}\sum_{i=1}^{n} K_i(x_0, T - t_0; u_1^*, \ldots, u_n^*) &= \\ &= \max_{u_1, \ldots, u_n} \sum_{i=1}^{n} K_i(x_0, T - t_0; u_1, \ldots, u_n) = \\ &= \sum_{i=1}^{n} \int_{t_0}^{T} h_i(x^*(t))dt = v(N; x_0, T - t_0),\end{aligned}$$

$x^*(t)$ – cooperative trajectory.

Characteristic Function in $\Gamma(x_0, T - t_0)$.

$v(S; x_0, T - t_0), \quad S \subset N$,

superadditivity: $v(S_1 \cup S_2; x_0, T - t_0) \geq v(S_1; x_0, T - t_0) + v(S_2; x_0, T - t_0)$,
$S_1 \cap S_2 = \emptyset$.

There are different ways on how to define c. f.

a. Classical: $v(S; x_0, T - t_0) = Val\Gamma_{S,N\setminus S}(x_0, T - t_0)$, where $\Gamma_{S,N\setminus S}(x_0, T - t_0)$ is a zero-sum game played upon the structure of game $\Gamma(x_0, T - t_0)$ between S as player 1 and $N \setminus S$ as player 2.

b. $v(S; x_0, T - t_0) = \sum_{i\in S} K_i(x_0, T - t_0; \bar{u}_S, \bar{u}_{N\setminus S})$, where $(\bar{u}_S, \bar{u}_{N\setminus S})$ is some giver NE in $\Gamma'_{S,N\setminus S}$ played as non zero-sum game over the structure of $\Gamma(x_0, T - t_0)$ between two players: coalition S as player 1 and $N \setminus S$ as player 2

c. $v(S; x_0, T - t_0) = \max_{u_S=\{u_i, i\in S\}} \sum_{i\in S} K_i(x_0, T - t_0; \bar{u}||u_S)$, where $\bar{u} = (\bar{u}_1, \ldots, \bar{u}_n)$ is some fixed NE in $\Gamma(x_0, T - t_0)$.

L. Petrosjan, G. Zaccour Time-consistent Shapley value allocation of pollution cost reduction // Journal of Economics Dunamics & Control, 27 (2003), pp. 381-398.

Let $E(x_0, T - t_0)$ be the imputation set in $\Gamma(x_0, T - t_0)$:

$$E(x_0, T-t_0) = \{\xi = (\xi_i) : \sum_{i=1}^{n} \xi_i = v(N; x_0, T-t_0), \xi_i \geq v(\{i\}; x_0, T-t_0), i \in N\}.$$

Denote by $C^{t-t_0}(x_0), t \in [t_0, T]$ reachable set of the (2).
For each $y \in C^{t-t_0}(x_0)$ consider a subgame $\Gamma(y, T-t)$ of the game $\Gamma(x_0, T-t_0)$, with corresponding c. f. $v(S; y, T-t)$ and set of imputations $E(y, T-t)$.

Definition. A point-to-set mapping $C(y, T-t) \subset E(y, T-t)$ defined for all $y \in C^{t-t_0}(x_0), t \in [t_0, T]$ is call *solution concept* (SC) in the family of subgames $\Gamma(y, T-t)$.

In special cases $C(y, T-t)$ may be a core, NM-solution, Shapley value, nucleous etc.

What happens when the game develops along the cooperative trajectory $x^*(t)$? We pass through current subgame $\Gamma(x^*(t), T-t)$, wilingly or not updating the current SC $\leftrightarrow C(x^*(t), T-t)$.

Imputation Distribution Procedure (IDP).

Let $\bar{\xi} \in C(x_0, T - t_0)$ and $\beta_i(t), i \in N, t \in [t_0, T]$ satisfies the condition

$$\bar{\xi} = \int_{t_0}^{T} \beta_i(t)dt, \quad i \in N, \quad \beta_i \geq 0.$$

$\beta_i(t)$ is called IDP.

Define

$$\bar{\xi}(\theta) = \int_{t_0}^{\theta} \beta_i(t)dt, \quad i \in N, \quad \beta_i \geq 0.$$

Definition. The SC $C(x^*(t), T - t), t \in [t_0, T]$ is called *time-consistent* (TC) if there exist such IDP $\beta(t) = \{\beta_i(t)\}$ that

$$\bar{\xi} - \bar{\xi}(\theta) \in C(x^*(\theta), T - \theta)$$

for all $\theta \in [t_0, T]$.

Definition. The SC $C(x^*(t), T - t), t \in [t_0, T]$ is called *strongly time-consistent* (STC) if there exist such IDP $\beta(t) = \{\beta_i(t)\}$ that

$$\bar{\xi}(\theta) \oplus \bar{C}(x^*(\theta), T - \theta) \subset C(x_0, T - t_0)$$

for all $\theta \in [t_0, T]$. Here $\bar{\xi} \oplus A$ means the set of all possible vectors $\bar{\xi} + \eta$ for all $\eta \in A$.

Consider $C(x^*(t), T-t)$ along $x^*(t), t \in [t_0, T]$. Suppose we can construct a differentiable selector $\xi^t \in C(x^*(t), T-t)$, then we can easily get for $\beta(t)$ the following formula

$$\bar{\xi} = \bar{\xi}(\theta) + \xi^t \quad \rightarrow \quad \bar{\xi} = \int_{t_0}^{\theta} \beta_i(t)dt + \xi^t$$

$$\beta_i(t) = -\frac{d}{dt}\xi^t$$

If ξ^t can be chosen as monotonic nonincreasing (which is very possible since $h_i > 0$, then $\beta_i \geq 0$, and SC is TC.
If the case (for instance) $C(y, T-t)$ is a Shapley value, we get

$$\beta_i(t) = -\sum_{S \subset N, S \ni i} \frac{(n-s)!(s-1)!}{n!} \left[\frac{d}{dt} v(x^*(t), T-t; S) - \right.$$

$$\left. -\frac{d}{dt} v(x^*(t), T-t; S \setminus \{i\}) \right]$$

and we need only differentiability of the value function (c. f.) $v(x, T-t; S)$.

Strategically Supported Cooperation

Continuous time case.

Consider n-person differential game $\Gamma(x_0, T - t_0)$ with prescribed duration and independent motions on the time interval $[t_0, T]$. Motion equations have the form:

$$\dot{x}_i = f_i(x_i, u_i), \quad u_i \in U_i \subset R^{\ell}, x_i \in R^n,$$
$$x_i(t_0) = x_i^0, \quad i = 1, \dots, n. \tag{2}$$

It is assumed that the system of differential equations (2) satisfies all conditions necessary for the existence, prolongability and uniqueness of the solution for any n-tuple of measurable controls $u_1(t), \dots, u_n(t)$.

The payoff of player i is defined as:

$$H_i(x_0, T - t_0; u_1(\cdot), \dots, u_n(\cdot)) = \int_{t_0}^{T} h_i(x(\tau))d\tau,$$

where $h_i(x)$ is a continuous function and $x(\tau) = \{x_1(\tau), \dots, x_n(\tau)\}$ is the solution of (2) when open-loop controls $u_1(t), \dots, u_n(t)$ are used and $x(t_0) = \{x_1(t_0), \dots, x_n(t_0)\} = \{x_1^0, \dots, x_n^0\}$.

Suppose that there exist an n-tuple of open-loop controls $\bar{u}(t) = (\bar{u}_1(t), \ldots, \bar{u}_n(t))$ and the trajectory $\bar{x}(t)$, $t \in [t_0, T]$, such that

$$\max_{u_1(t),\ldots,u_n(t)} \sum_{i=1}^{n} H_i(x_0, T - t_0; u_1(t), \ldots, u_n(t)) =$$
$$= \sum_{i=1}^{n} H_i(x_0, T - t_0; \bar{u}_1(t), \ldots, \bar{u}_n(t)) = \sum_{i=1}^{n} \int_{t_0}^{T} h_i(\bar{x}_i(\tau))d\tau \quad (3)$$

The trajectory $\bar{x}(t) = (\bar{x}_1(t), \ldots, \bar{x}_n(t))$ satisfying (3) we shall call "optimal cooperative trajectory".
Let $N = \{1, \ldots, n\}$ be the set of players. Define in $\Gamma(x_0, T - t_0)$ characteristic function in a classical way:

$$V(x_0, T - t_0; N) = \sum_{i=1}^{n} \int_{t_0}^{T} h_i(\bar{x}_i(\tau))d\tau,$$
$$V(x_0, T - t_0; \emptyset) = 0,$$
$$V(x_0, T - t_0; S) = Val\, \Gamma_{S,N\setminus S}(x_0, T - t_0), \quad (4)$$

where $Val\, \Gamma_{S,N\setminus S}(x_0, T - t_0)$ is a value of zero-sum game played between coalition S acting as first player and coalition $N \setminus S$ acting as player 2, with payoff of player S equal to:

$$\sum_{i\in S} H_i(x_0, T - t_0; u_1(\cdot), \ldots, u_n(\cdot)).$$

Define $L(x_0, T - t_0)$ as imputation set in the game $\Gamma(x_0, T - t_0)$ (see Neumann and Morgenstern (1947)):

$$L(x_0, T - t_0) = \{\alpha = (\alpha_1, \dots, \alpha_n) : \\ \left. \alpha_i \geqslant V(x_0, T - t_0; \{i\}), \quad \sum_{i \in N} \alpha_i = V(x_0, T - t_0; N) \right\}. \quad (5)$$

Regularized game $\Gamma_\alpha(x_0, T - t_0)$. For every $\alpha \in L(x_0, T - t_0)$ define the noncooperative game $\Gamma_\alpha(x_0, T - t_0)$, which differs from the game $\Gamma(x_0, T - t_0)$ only by payoffs defined along optimal cooperative trajectory $\bar{x}(\tau)$, $\tau \in [t_0, T]$.
Let $\alpha \in L(x_0, T - t_0)$. Define the imputation distribution procedure (IDP) (see Petrosjan (1993)) as function $\beta(\tau) = (\beta_1(\tau), \dots, \beta_n(\tau))$, $\tau \in [t_0, T]$ such that

$$\alpha_i = \int_{t_0}^{T} \beta_i(\tau) d\tau. \quad (6)$$

Denote by $H_i^\alpha(x_0, T - t_0; u_1(\cdot), \dots, u_n(\cdot))$ the payoff function in the game $\Gamma_\alpha(x_0, T - t_0)$ and by $x(\tau)$ the corresponding trajectory, then

$$H_i^\alpha(x_0, T - t_0; u_1(\cdot), \dots, u_n(\cdot)) = H_i(x_0, T - t_0; u_1(\cdot), \dots, u_n(\cdot))$$

if there does not exist such $t \in [t_0, T]$ that $x(\tau) = \bar{x}(\tau)$ for $\tau \in (t_0, t]$.

Let $t = \sup\{t' : x(\tau) = \bar{x}(\tau), \tau \in [t_0, t']\}$ and $t > t_0$, then

$$H_i^{\alpha}(x_0, T - t_0; u_1(\cdot), \dots, u_n(\cdot)) =$$
$$= \int_{t_0}^{t} \beta_i(\tau) d\tau + H_i(\bar{x}(t), T - t; u_1(\cdot), \dots, u_n(\cdot)) =$$
$$= \int_{t_0}^{t} \beta_i(\tau) d\tau + \int_{t}^{T} h_i(x(\tau)) d\tau.$$

In a special case, when $x(\tau) = \bar{x}(\tau)$, $\tau \in [t_0, T]$ (if $x(\tau)$ is an optimal cooperative trajectory in the sense of Eq. (3)), we have

$$H_i^{\alpha}(x_0, T - t_0; \bar{u}_1(\cdot), \dots, \bar{u}_n(\cdot)) = \int_{t_0}^{T} \beta_i(\tau) d\tau = \alpha_i.$$

By the definition of payoff function in the game $\Gamma_{\alpha}(x_0, T - t_0)$we get that the payoffs along the optimal trajectory are equal to the components of the imputation $\alpha = (\alpha_1, \dots, \alpha_n)$.
Consider the current subgames (see Neumann and Morgenstern (1947)) — $\Gamma(\bar{x}(t), T - t)$ along $\bar{x}(t)$ and current imputation sets $L(\bar{x}(t), T - t)$. Let $\alpha(t) \in L(\bar{x}(t), T - t)$. Suppose that $\alpha(t)$ can be selected as differentiable function of t, $t \in [t_0, T]$.

Definition 1. The game $\Gamma_\alpha(x_0, T - t_0)$ is called regularization of the game $\Gamma(x_0, T - t_0)$ (α-regularization) if the IDP β is defined in such a way that

$$\alpha_i(t) = \int_t^T \beta_i(\tau) d\tau$$

or

$$\beta_i(t) = -\alpha_i'(t). \tag{7}$$

Theorem 1. In the regularization of the game $\Gamma_\alpha(x_0, T - t_0)$ for every $\varepsilon > 0$ there exist an ε-Nash equilibrium (Nash (1951)) with payoffs $\alpha = (\alpha_1, \ldots, \alpha_i, \ldots, \alpha_n)$.

Proof. The proof is based on actual constraction of the ε-Nash equilibrium in piecewise open-loop (POL) strategies with memory.
Remind the definition of POL strategies with memory in differential game. Denote by $\hat{x}(t)$ any admissible trajectory of the system (2) on the time interval $[t_0, t]$, $t \in [t_0, T]$.
The strategy $u_i(\cdot)$ of player i is called POL if it consists from the pair (a, σ), where σ is a partition of time interval $[t_0, T]$, $t_0 < t_1 < \ldots < t_l = T$ ($t_{k+1} - t_k = \delta > 0$), and a mapping a which corresponds to each point $(\hat{x}(t_k), t_k)$, $t_k \in \sigma$ an open-loop control $u_i(t)$, $t \in [t_k, t_{k+1})$.

Consider a family of associated with $\Gamma(x, T-t)$, but not with $\Gamma_\alpha(x, T-t)$ zero-sum games $\Gamma_{\{i\},N\setminus\{i\}}(x, T-t)$ from the initial position x and duration $T-t$ between the coalition S consisting from a single player i and the coalition $N\setminus\{i\}$ with player's i payoff equal to

$$H_i(x, T-t; u_1(\cdot)\ldots, u_n(\cdot)).$$

The payoff of player $N\setminus\{i\}$ in $\Gamma_{\{i\},N\setminus\{i\}}(x, T-t)$ equals to $(-H_i)$. Let $\hat{u}(x, t; \cdot)$ be the ε-optimal POL strategy of player $N\setminus\{i\}$ in $\Gamma_{\{i\},N\setminus\{i\}}(x, T-t)$. Note, that $\hat{u}(x, t; \cdot) = \{\hat{u}_j(x, t; \cdot)\}$, $j \in N\setminus\{i\}$.
Let $\hat{x}(t) = \{\hat{x}_1(t), \ldots, \hat{x}_n(t)\}$ be the segment of an admissible trajectory of (2) defined on the time interval $[t_0, t]$, $t \in [t_0, T]$. For each $i \in \{1, \ldots, n\}$ define $\bar{t}(i) = \sup\{t_i : \hat{x}_i(t_i) = \hat{\hat{x}}_i(t_i)\}$ and $\bar{\bar{t}}(j) = \min_i \bar{t}(i) = \bar{t}(j)$. $\bar{\bar{t}}(j)$ lies in one of the intervals $[t_k, t_{k+1})$, $k = 0, 1, \ldots, l-1$. Thus, $\bar{t}(i) - t_0$ is the length of the time interval starting from t_0 on which $x_i(t)$ coincides with $\bar{x}_i(t)$ — the i-th component of the cooperative trajectory $\bar{x}(t)$. And $\bar{\bar{t}}(j) - t_0$ is the length of the time interval starting from t_0 on which $x(t)$ coincides with cooperative trajectory $\bar{x}(t)$.

Define the following strategies of player $i \in N$.

$$u_i^*(\cdot) = \begin{cases} \bar{u}_i(t) & \text{for } (\hat{x}(t_k), t_k) \text{ on the optimal cooperative} \\ & \text{trajectory } \bar{x}(t) \;\; (\hat{x}(\tau) = \bar{x}(\tau), \tau \in [t_0, t_k]); \\ \hat{u}_i(\hat{x}(t_{k+1}), t_{k+1}; \cdot) & i\text{-th component of the } \varepsilon/2\text{-optimal POL} \\ & \text{strategy of player } N \setminus \{j\} \text{ in the game} \\ & \Gamma_{\{j\}, N\setminus\{j\}}(x(t_{k+1}), T - t_{k+1}), \text{ if } t_k \leqslant \bar{\bar{t}}(j) < t_{k+1}; \\ \text{arbitrary} & \text{for all other positions.} \end{cases}$$

Show that $u^*(\cdot) = (u_1^*(\cdot), \ldots, u_n^*(\cdot))$ is ε-Nash equilibrium in $\Gamma_\alpha(x_0, T - t_0)$. The following equality holds

$$H_i(x_0, T - t_0; u^*(\cdot)) = H_i(x_0, T - t_0; u_1^*(\cdot), \ldots, u_n^*(\cdot)) = \int_{t_0}^{T} \beta_i(t)dt = \alpha_i. \quad (8)$$

Consider the n-tuple $(u^*(\cdot)||u_i(\cdot))$ where player i changes his strategy $u_i^*(\cdot)$ on $u_i(\cdot)$.

We have to show that

$$H_i(x_0, T - t_0; u^*(\cdot)) \geqslant H_i(x_0, T - t_0; u^*(\cdot)||u_i(\cdot)) - \varepsilon. \quad (9)$$

for all $i \in N$ and all POL $u_i(\cdot)$ of player i.

It is easy to see that when the n-tuple $u^*(\cdot)$ is played the game develops along the optimal trajectory $\bar{x}(t)$. If in $(u^*(\cdot)||u_i(\cdot))$ the trajectory $\bar{x}(t)$ is also realized then (9) will be equality and thus true.

Suppose now that in $(u^*(\cdot)||u_i(\cdot))$ the trajectory $x(t)$ different form $\bar{x}(t)$ is realized. Then let

$$\bar{t} = \inf\{t : \bar{x}(t) \neq x(t)\}.$$

and $\bar{t} \in [t_{k-1}, t_k)$. Since the motion of players are independent we get $\bar{x}_m(t_k) = x_m(t_k)$ for $m \in N \setminus \{i\}$ and $\bar{x}_i(t_k) \neq x_i(t_k)$ (but $\bar{x}_j(t_{k-1}) = x_j(t_{k-1})$ for $j \in N$). Then from the definition of $u^*(\cdot)$ it follows that the players $m \in N \setminus \{i\}$ will use their strategies $\hat{u}_m(\hat{x}(t_k), t_k; \cdot)$ which are $\frac{\varepsilon}{2}$-optimal in a zero-sum game $\Gamma_{\{i\}, N\setminus\{i\}}(x(t_k), T - t_k)$ against the player i which deviates from the optimal trajectory on a time interval $[t_{k-1}, t_k)$.
If the players from the set $N \setminus \{i\}$ will use their strategies $\hat{u}_m(\hat{x}(t_k), t_k; \cdot)$, player i starting from position $x(t_k), t_k$ will get not more than

$$V(x(t_k), T - t_k; \{i\}) + \frac{\varepsilon}{2},$$

where $V(x(t_k), T - t_k; \{i\})$ is the value of the game $\Gamma_{\{i\}, N\setminus\{i\}}(x(t_k), T - t_k)$. Then the total payoff of player i in $\Gamma_\alpha(x_0, T - t_0)$ when the n-tuple of strategies $(u^*(\cdot)||u_i(\cdot))$ is played cannot exceed the amount

$$\int_{t_0}^{t_{k-1}} \beta_i(\tau)d\tau + V(x(t_k), t_k; \{i\}) + \frac{\varepsilon}{2} + \int_{t_{k-1}}^{t_k} h_i(x_i(\tau))d\tau. \qquad (10)$$

But the payoff of player i when the n-tuple $u^*(\cdot)$ is played is equal to

$$\alpha_i = \int_{t_0}^{T} \beta_i(\tau)d\tau = \int_{t_0}^{t_{k-1}} \beta_i(\tau)d\tau + \int_{t_{k-1}}^{T} \beta_i(\tau)d\tau = \int_{t_0}^{t_{k-1}} \beta_i(\tau)d\tau + \alpha_i(t_{k-1}). \tag{11}$$

By the definition of IDP (see (6), (7)), $\alpha_i(t_{k-1}) \in L(\bar{x}(t_{k-1}), T - t_{k-1})$,

$$\int_{t_{k-1}}^{T} \beta_i(\tau)d\tau = \alpha_i(t_{k-1}) \geqslant V(\bar{x}(t_{k-1}), T - t_{k-1}; \{i\}). \tag{12}$$

From the continuity of the function V and continuity of the trajectory $x(t)$ by appropriate choice of $\delta > 0$ ($t_{k+1} - t_k = \delta$) the following inequalities can be guaranteed:

$$|V(\bar{x}(t_{k-1}), T - t_{k-1}; \{i\}) - V(x(t_k), T - t_k; \{i\})| < \frac{\varepsilon}{4},$$

$$\int_{t_{k-1}}^{T} \beta_i(\tau)d\tau = \alpha_i(t_{k-1}) \geqslant V(x(t_k), T - t_k; \{i\}) - \frac{\varepsilon}{4}.$$

Compare $\alpha_i(t_{k-1})$ and $V(x(t_k), t_k; \{i\}) + \frac{\varepsilon}{2} + \int_{t_{k-1}}^{t_k} h_i(x_i(\tau))d\tau$. By choosing $\delta = t_{k+1} - t_k$ sufficiently small one can achieve that the integral $\int_{t_{k-1}}^{t_k} h_i(x_i(\tau))d\tau$ will be also small (less than $\varepsilon/4$).

Adding to both sides of (12) the amount $\int_{t_0}^{t_{k-1}} \beta_i(\tau)d\tau$ and using the previous inequality we get

$$\alpha_i = \int_{t_0}^{t_{k-1}} \beta_i(\tau)d\tau + \alpha_i(t_{k-1}) \geqslant \int_{t_0}^{t_{k-1}} \beta_i(\tau)d\tau + V(\bar{x}(t_{k-1}), T - t_{k-1}; \{i\}) \geqslant$$
$$\geqslant \int_{t_0}^{t_{k-1}} \beta_i(\tau)d\tau + V(x(t_k), T - t_k; \{i\}) - \frac{\varepsilon}{4}$$
$$\geqslant \int_{t_0}^{t_{k-1}} \beta_i(\tau)d\tau + V(x(t_k), T - t_k; \{i\}) - \frac{\varepsilon}{4} + \int_{t_{k-1}}^{t_k} h_i(\tau)d\tau - \frac{\varepsilon}{4}$$
$$\geqslant \int_{t_0}^{t_{k-1}} \beta_i(\tau)d\tau + V(x(t_k), T - t_k; \{i\}) + \int_{t_{k-1}}^{t_k} h_i(\tau)d\tau - \frac{\varepsilon}{2}$$
$$\geqslant \int_{t_0}^{t_{k-1}} \beta_i(\tau)d\tau + V(x(t_k), T - t_k; \{i\}) + \int_{t_{k-1}}^{t_k} h_i(\tau)d\tau +$$
$$+\frac{\varepsilon}{2} - \frac{\varepsilon}{2} - \frac{\varepsilon}{2}. \tag{13}$$

Here first four addends in the right part of the inequality constitute the upper bound of player i payoff when $(u^*(\cdot)||u_i^*(\cdot))$ is played. But α_i is the payoff of player i when $u^*(\cdot)$ is played, and we get

$$\begin{aligned} H_i(x_0, T - t_0; u^*(\cdot)) = \alpha_i \geqslant \\ \geqslant \int_{t_0}^{t_{k-1}} \beta_i(\tau)d\tau + V(x(t_k), T - t_k; \{i\}) + \int_{t_{k-1}}^{t_k} h_i(\tau)d\tau + \frac{\varepsilon}{2} - \varepsilon \geqslant \\ \geqslant H_i(x_0, T - t_0; u^*(\cdot)||u_i(\cdot)) - \varepsilon \end{aligned} \tag{14}$$

and we get (9). The theorem is proved. □

This means that the cooperative solution (any imputation) can be strategically supported in a regularized game $\Gamma_\alpha(x_0, T - t_0)$ (realized in a specially constructed Nash equilibrium) by the Nash equilibrium $u^*(\cdot)$ defined in the Theorem 1.

Discrete time case.

In what follows as basic model we shall consider the game in extensive form with perfect information.

Definition 2. A game tree is a finite oriented treelike graph K with the root x_0. We shall use the following notations. Let x be some vertex (position). We denote by $K(x)$ a subtree K with the root in x. We denote by $Z(x)$ immediate successors of x. The vertices y, directly following after x, are called alternatives in x ($y \in Z(x)$). The player who makes a decision in x (who selects the next alternative position in x), will be denoted by $i(x)$. The choice of player $i(x)$ in position x will be denoted by $\bar{x} \in Z(x)$.

Let $N = \{1, \dots, n\}$ — be the set of all players in the game.

Definition 3. A game in extensive form with perfect information (see Kuhn (1953)) $G(x_0)$ is a graph tree $K(x_0)$, with the following additional properties:

- The set of vertices (positions) is split up into $n+1$ subsets $P_1, P_2, \dots, P_{n+1}$, which form a partition of the set of all vertices of the graph tree K. The vertices (positions) $x \in P_i$ are called players i personal positions, $i = 1, \dots, n$; vertices (positions) $x \in P_{n+1}$ are called terminal positions.
- In each final vertex (position) the system of real numbers $h(w) = (h_1(w), \dots, h_n(w))$, $w \in P_{n+1}$, $h_i(w) \geqslant 0$, $i = 1, \dots, n$ is defined. Where $h_i(w)$ is the payoff of player i in the final vertex (position).

Definition 4. A strategy of player i is a mapping $U_i(\cdot)$, which associate to each position $x \in P_i$ a unique alternative $y \in Z(x)$.
As in the previous case denote by $H_i(x; u_1(\cdot), \dots, u_n(\cdot))$ the payoff function od player $i \in N$ in the subgame $G(x)$ starting from the position x.

$$H_i(x; u_1(\cdot), \dots, u_n(\cdot)) = h_i(x'_l)$$

where $x'_l \in P_{n+1}$ is the last vertex (position) in the path $x = (x'_1, x'_2, \dots, x'_l)$ realized in the subgame $G(x)$, when the n-tuple of strategies $(u_1(\cdot), \dots, u_n(\cdot))$ is played.
Denote by $\bar{u}(\cdot) = (\bar{u}_1(\cdot), \dots, \bar{u}_n(\cdot))$ the n-tuple of strategies and the trajectory (path) $\bar{x} = (\bar{x}_0, \bar{x}_1, \dots, \bar{x}_m)$, $\bar{x}_m \in P_{n+1}$ such that

$$\max_{u_1(\cdot), \dots, u_n(\cdot)} \sum_{i=1}^{n} H_i(x_0; u_1(\cdot), \dots, u_n(\cdot)) =$$

$$= \sum_{i=1}^{n} H_i(x_0; \bar{u}_1(\cdot), \dots, \bar{u}_n(\cdot)) = \sum_{i=1}^{n} h_i(\bar{x}_m). \qquad (15)$$

The path $\bar{x} = (\bar{x}_0, \ldots, \bar{x}_m)$ satisfying Eq. (15) we shall call "optimal cooperative trajectory".
Define in $G(x_0)$ characteristic function in a classical way

$$
\begin{aligned}
&V(x_0; N) = \sum_{i=1}^{n} h_i(\bar{x}_m),\\
&V(x_0; \emptyset) = 0,\\
&V(x_0; S) = Val\ \Gamma_{S, N \setminus S}(x_0),
\end{aligned}
$$

where $Val\ \Gamma_{S, N\setminus S}(x_0)$ is a value of zero-sum game played between coalition S acting as first player and coalition $N \setminus S$ acting as player 2, with payoff of player S equal to

$$
\sum_{i \in S} H_i(x_0; u_1(\cdot), \ldots, u_n(\cdot)).
$$

Define $L(x_0)$ as imputation set in the game $G(x_0)$.

$$
L(x_0) = \left\{ \alpha = (\alpha_1, \ldots, \alpha_n) : \alpha_i \geqslant V(x_0; \{i\}), \quad \sum_{i \in N} \alpha_i = V(x_0; N) \right\}.
$$

Regularized game $G_\alpha(x_0)$. For every $\alpha \in L(x_0)$ define the noncooperative game $G_\alpha(x_0)$, which differs from the game $G(x_0)$ only by payoffs defined along optimal cooperative path $\bar{x} = (\bar{x}_0, \ldots, \bar{x}_m)$. Let $\alpha \in L(x_0)$. Define the imputation distribution procedure (IDP) as function $\beta_k = (\beta_1(k), \ldots, \beta_n(k))$, $k = 0, 1, \ldots, m$ such that

$$\alpha_i = \sum_{k=0}^{m} \beta_i(k). \tag{16}$$

Define by $H_i^\alpha(x_0; u_1(\cdot), \ldots, u_n(\cdot))$ the payoff function in the game $G_\alpha(x_0)$ and by $\bar{x} = \{\bar{x}_0, \ldots, \bar{x}_m\}$ the cooperative path

$$H_i^\alpha(x_0; u_1(\cdot), \ldots, u_n(\cdot)) = H_i(x_0; u_1(\cdot), \ldots, u_n(\cdot))$$

for all $u_1(\cdot), \ldots, u_n(\cdot)$ such that the path $x = \{x_0, \ldots, x_m\}$ differs from $\bar{x} = \{\bar{x}_0, \ldots, \bar{x}_m\}$, and

$$H_i^\alpha(x_0; \bar{u}_1(\cdot), \ldots, \bar{u}_n(\cdot)) = \alpha_i.$$

By the definition of the payoff function in the game $G_\alpha(x_0)$ we get that the payoffs along the optimal cooperative trajectory are equal to the components of the imputation $\alpha = (\alpha_1, \ldots, \alpha_n)$.

Consider current subgames $G(\bar{x}_k)$ along the optimal path $\bar{x}$ and current imputation sets $L(\bar{x}_k)$. Let $\alpha^k \in L(\bar{x}_k)$.
Definition 5. The game $G_\alpha(x_0)$ is called regularization of the game $G(x_0)$ (α-regularization) if the IDP β is defined in such a way that

$$\alpha_i^k = \sum_{j=k}^{m} \beta_i(j)$$

or $\beta_i(k) = \alpha_i^k - \alpha_i^{k+1}$, $i \in N$, $k = 0, 1, \ldots, m-1$, $\beta_i(m) = \alpha_i^m$, $\alpha_i^0 = \alpha_i$.
Theorem 2. In the regularization of the game $G_\alpha(x_0)$ there exist a Nash equilibrium with payoffs $\alpha = (\alpha_1, \ldots, \alpha_n)$.
Proof. Along the cooperative path we have

$$\alpha_i^k \geqslant V(\bar{x}_k; \{i\}), \quad i \in N, k = 0, 1, \ldots, m.$$

since $\alpha^k = (\alpha_1^k, \ldots, \alpha_n^k) \in L(\bar{x}_k)$ is an imputation in $G(\bar{x}_k)$ (note that here $V(\bar{x}_k; \{i\})$ is computed in the subgame $G(\bar{x}_k)$ but not $G_\alpha(\bar{x}_k)$). In the same time

$$\alpha_i^k = \sum_{j=k}^{m} \beta_i(j)$$

and we get

$$\sum_{j=k}^{m} \beta_i(j) \geqslant V(\bar{x}_k; \{i\}), \quad i \in N, k = 0, 1, \dots, m. \tag{17}$$

But $\sum_{j=k}^{m} \beta_i(j)$ is the payoff of player i in the subgame $G_\alpha(\bar{x}_k)$ along the cooperative path, and from (17) using the arguments similar to those in the proof of Theorem 1 one can construct the Nash equilibrium with payoffs $\alpha = (\alpha_1, \dots, \alpha_n)$ and resulting cooperative path $\bar{x} = (\bar{x}_0, \dots, \bar{x}_m)$.

Example. In this example as an imputation we shall consider Shapley value [Shapley (1953)]. Using the proposed regularization of the game we shall see that there exist a Nash equilibrium with payoffs equal to the components of the Shapley value.

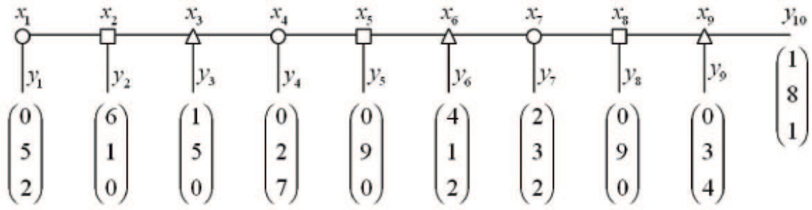


Fig. 1. Game $G(x_0)$

In the game $G(x_0)$, $N = \{1, 2, 3\}$, $P_1 = \{x_1, x_4, x_7\}$, $P_2 = \{x_2, x_5, x_8\}$, $P_3 = \{x_3, x_6, x_9\}$, $P_4 = \{y_1, y_2, y_3, y_4, y_5, y_6, y_7, y_8, y_9, y_{10}\}$. $h(y_1) = (0, 5, 2)$, $h(y_2) = (6, 1, 0)$, $h(y_3) = (1, 5, 0)$, $h(y_4) = (0, 2, 7)$, $h(y_5) = (0, 9, 0)$, $h(y_6) = (4, 1, 2)$, $h(y_7) = (2, 3, 2)$, $h(y_8) = (0, 9, 0)$, $h(y_9) = (0, 3, 4)$, $h(y_{10}) = (1, 8, 1)$. The cooperative path is $\bar{x} = \{\bar{x}_1, \bar{x}_2, \bar{x}_3, \bar{x}_4, \bar{x}_5, \bar{x}_6, \bar{x}_7, \bar{x}_8, \bar{x}_9, \bar{y}_{10}\}$.

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	y_{10}
$V(x;\{1\})$	0	0	0	0	0	2	2	0	0	1
$V(x;\{2\})$	2	2	2	2	9	1	3	9	3	8
$V(x;\{3\})$	0	0	0	0	0	2	0	0	4	1
$V(x;\{1,2\})$	7	7	6	9	9	5	9	9	3	9
$V(x;\{2,3\})$	7	9	9	9	9	5	5	9	9	9
$V(x;\{1,3\})$	6	6	6	7	0	6	4	0	4	2
$V(x;\{1,2,3\})$	10	10	10	10	10	10	10	10	10	10
$Sh(x;\{1\})$	$\frac{17}{6}$	$\frac{13}{6}$	$\frac{12}{6}$	$\frac{16}{6}$	$\frac{2}{6}$	$\frac{22}{6}$	$\frac{24}{6}$	$\frac{2}{6}$	$\frac{2}{6}$	1
$Sh(x;\{2\})$	$\frac{26}{6}$	$\frac{28}{6}$	$\frac{27}{6}$	$\frac{28}{6}$	$\frac{56}{6}$	$\frac{16}{6}$	$\frac{30}{6}$	$\frac{56}{6}$	$\frac{26}{6}$	8
$Sh(x;\{3\})$	$\frac{17}{6}$	$\frac{19}{6}$	$\frac{21}{6}$	$\frac{16}{6}$	$\frac{2}{6}$	$\frac{22}{6}$	$\frac{6}{6}$	$\frac{2}{6}$	$\frac{32}{6}$	1

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	y_{10}
$\beta_1(j)$	$\frac{2}{6}$	$\frac{4}{6}$	$\frac{1}{6}$	$\frac{14}{6}$	$-\frac{28}{6}$	$-\frac{2}{6}$	$\frac{22}{6}$	0	$-\frac{4}{6}$	1
$\beta_2(j)$	$-\frac{1}{6}$	$-\frac{2}{6}$	$\frac{1}{6}$	$-\frac{28}{6}$	$\frac{40}{6}$	$-\frac{14}{6}$	$-\frac{28}{6}$	$\frac{30}{6}$	$-\frac{22}{6}$	8
$\beta_3(j)$	$-\frac{1}{6}$	$-\frac{2}{6}$	$-\frac{2}{6}$	$\frac{14}{6}$	$-\frac{20}{6}$	$\frac{16}{6}$	$\frac{4}{6}$	$-\frac{30}{6}$	$\frac{26}{6}$	1

It can be easily seen that the inequality (17)

$$\sum_{j=k}^{m} \beta_i(j) \geqslant V(\bar{x}_k; \{i\})$$

for $i \in N$ holds in this case.

References

Neumann J., Morgenstern O. (1947). "Theory of Games and Economic Behavior Princeton.

Nash J. (1951). "Non-cooperative games Ann. Mathematics, 54, pp. 286–295.

Petrosjan L. A. (1993). "Differential Games of Pursuit World Scientific, Singapore.

Petrosyan L. A., Yeung D. W. K. (2006) Cooperative Stochastic Differential Games. Springer, 242 p.

Shapley L.S. (1953). "A Value for *n*-Person Games Contributions to the Theory of Games (in H.W. Kuhn and A.W. Tucker (eds.), Princeton, Princeton University Press, pp. 307–315.

Petrosyan L.A., Zenkevich N.A. (2009) Conditions for Sustainable Cooperation. Contributions to Game Theory and Management, SPb, GSOM, Vol.2, pp.344–355.

Petrosyan L.A., Danilov N.N. (1979) Ustojchivost reshenij neantagonisticheskih differencialnyh igr s transferabelnymi vyigryshami. Vestnik Leningradskogo universiteta, pp. 52–59 (in russian).

Petrosyan L.A. (1977) Ustojchivost reshenij differencialnyh igr so mnogimi uchastnikami. Vestnik Leningradskogo universiteta, pp. 46–52 (in russian).